Modelling the constraints on the global pharmaceutical industry

Geoffrey D Hobbs
8a Alma Road
Reigate
Surrey, RH2 0DA, UK

Bryan R Deane
SmithKline Beecham Pharmaceuticals
SB House
Great West Road
Brentford
Middlesex, TW8 9BD, UK

Abstract

Advances in all fields of medical technology have driven rapidly growing expectations of medical care over the last half century. The rate of growth of this demand for health-care has consistently exceeded GDP growth and, in many countries, the health-care bill has been absorbing an ever greater proportion of government spending. Governments throughout the world are increasingly concerned that this growth is unsustainable. The pharmaceutical industry, whose products account for about 5 to 10% of the health-care spend, has become an early target for a wide variety of cost containment measures. The industry's growth has been based on a cycle of growing sales from ever more effective new products, fuelling a substantial re-investment in high-risk, long-term research and development, leading to further advances and new product introductions. Historically the overall sales growth has been founded on both price-related and volume-related factors.

The purpose of this study was to explore the relevance of a System Dynamic modelling approach to understanding the potentially complex interactions between the pharmaceutical industry, the medical and related professions, and the regulators and paymasters who fund health-care. A prototype model has been constructed and used to create a variety of scenarios describing alternative futures, in which the regulators impose constraints on either price or volume increases. Not surprisingly, the developing dynamics vary, depending on the type and severity of the constraints and the industry's responses to them. For example, under price control, increased investment in marketing by the industry to promote volume growth might accelerate the reduction in the industry's profit margins, leading to a fall in the proportion of sales income devoted to R&D and a fall in the number of new products being developed. This scenario suggests a transformation of the industry towards a high volume, low margin, non-innovative "commodity" industry.

Experiments with the prototype suggest that the System Dynamics modelling approach *can* help to explore in an insightful way the potentially complex interactions of the various groups involved in health-care delivery. There is extremely broad scope for further development of the approach, which should ideally be targeted at specific issues within the system.

Modelling the Constraints on the Global Pharmaceutical Industry

1 Introduction

Over the last fifty years, dramatic advances have been made in all aspects of the practice of medicine, including surgical techniques, applied technology and supportive care as well as the introduction of new medicines. At the same time, the expectations of the educated populations of the world have risen dramatically. For any patient it is almost assumed that ever more heroic (and expensive) measures will be taken until either recovery is complete, or medical practice has exhausted its repertoire. This expectation has ensured that demand for health-care expenditure has grown faster than GDP in most developed countries for at least the last three decades. That populations are aging is both testimony to the medical success story, and further fuel for health-care demand.

Governments throughout the world are increasingly concerned that the escalating costs of health-care, claiming an ever greater share of government expenditure, are unsustainable. In the USA for example, it has been estimated that health-care accounted for the disposal of about 14% of GDP in 1992; about half of this was in the public sector, which was growing at about 13% pa (compared with about 3%pa GDP growth). Where, traditionally, decisions at the point of delivery (usually the doctor-patient interface) have been concerned with risk/benefit considerations, already cost is being factored in as another issue in the decision of whether or how to treat. Debate will undoubtedly continue as to what extent such decisions should be taken on purely medical grounds and to what extent it should be seen in a wider socio-economic context. One thing is certain: governments' ability to fund health-care is not unlimited. Undoubtedly, budgets for health-care spending will be controlled at some level.

Although the medicines bill accounts for only 5 to 10% of overall health-care expenditure in most countries, this has become an early target for cost containment measures. While the pharmaceutical industry is already highly regulated on the supply side, many of these regulations, (covering, for example, the quality of preclinical and clinical testing and manufacturing) tend to drive up industry's costs and therefore contribute to the inflating prices of newly-developed drugs. For several years there have been examples of direct controls on drug prices and/or drug company profitability, and governments are now increasingly investigating and implementing a variety of demand-side regulations (1) (including prescribing budgets, reference pricing systems, mandatory price controls or cuts, limited reimbursement lists, generic and therapeutic substitution and varying patient co-payment contributions).

For the global pharmaceutical industry, two aims are paramount: to contribute to real advances in medical treatment and to provide shareholders with an increasing (in real terms) flow of dividends. The pursuit of these goals has resulted in a rapid growth in sales, driven partly by volume growth and partly by rising prices. The opportunity for achieving sales growth through annual price increases is now, however, severely limited. In such a climate the only mechanism leading to rising average prices is the continual introduction of innovative new products at premium prices. While this may be critical for the sustained growth of individual companies, in a highly fragmented global industry, the impact of any one product introduction on global sales is minimal, at least in the first few years.

Other important drivers of global growth are volume-related. These can be classified as relating to population growth, to increasing need for medicines amongst certain groups within the population and to increasing access of the population to medicines. Population growth clearly leads to a potential increase in the volume of medicines consumed. Increasing need encompasses factors such as the increasing proportion of aging people, needing more medical care per person, and the emergence or spread of new or difficult-to-treat diseases, including AIDS. Increasing access depends upon economic

development allowing more people access to better medical care; while this is particularly relevant to developing countries, it is also an issue in, for example, the US, where bringing the millions of uninsured into a health-care system should in theory expand the potential market volume.

The growth aspirations of the pharmaceutical industry and the cost containment objectives of governments are fundamentally in conflict. Each might be expected to adopt measures aimed at securing success in its own objectives; each will respond to minimise the damage to its own aims. It would be surprising if the outcome were as either anticipated. At worst, the result might be seriously damaging - either to the industry or to the broader socio-economic aims of governments or to both.

Therefore the purpose of the study described here was to explore the relevance of a System Dynamics modelling approach to understanding the potentially complex interactions between the pharmaceutical industry, the medical and related professions, and the regulators and paymasters who fund health-care, much of it on behalf of governments. The model described here has been constructed using "iThink" software and has been used to create many scenarios describing alternative futures, some of which are discussed below. At this stage of the project, the model should be considered a prototype designed to examine whether further development of this type of approach might offer a powerful management learning tool (2,3) which would provide valuable insights into some of the issues which currently concern governments and, of necessity, the industry. For this reason, the scenarios described below should not be interpreted as quantitative forecasts or as the basis for policy recommendations.

2 Mapping the model's structure

2.1 The feedback loop structure

The aim, in developing the model, was to identify the most important mechanisms, described in terms of feedback loops, driving the growth of the pharmaceutical industry and then to superimpose the mechanisms that might limit its growth and bring the system into a dynamic equilibrium. The model, in its most basic form, is constructed from three positive (reinforcing) and two negative (balancing) feedback loops (Fig 1). The reinforcing loops are all associated

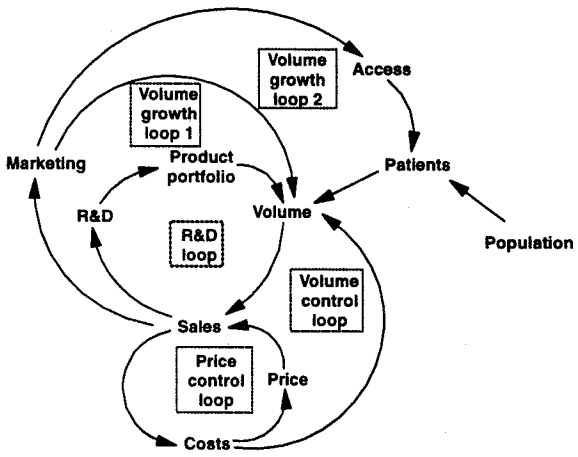

Figure 1. Model structure

with the pharmaceutical industry's growth aspirations. The balancing loops are associated with the efforts of the funding organisations to constrain growth. The growth of population is included as an exogenous driving force. Additional, subsidiary mechanisms have been added to increase the models realism.

The *R&D loop* represents the industry's core activity: the expenditure of part of its sales income on R&D in order to add to the portfolio of products it offers in the world's markets -a portfolio which generates a volume of drug usage which, in turn, creates further sales value and hence more R&D. Sales income is also used to finance marketing effort which in part promotes the more intensive use of existing drugs (eg by treating a wider range of illnesses -*volume growth loop 1*) and in part by increasing their accessibility (eg by expanding geographical market coverage - *volume growth loop 2*).

The two control loops represent the funders' responses to costs rising faster than they wish. They can attempt to limit growth by *price control*, by *volume control* or by a combination of both.

2.2 Industry driven growth

The core R&D loop is displayed, in outline only, in Figure 2. It is assumed that the industry allocates a percentage of its sales income to R&D and that this, when divided by the R&D expenditure incurred for every successful new product launched, creates a flow of new products which adds to the portfolio of existing products. In reality the R&D expenditure associated with a successful product is incurred over the 10-15 years prior to launch but this process is not captured in the models described here. The price at which a new product is launched is in general higher than that of existing products and therefore causes the average price to rise - a driving mechanism contributing to the growth of sales value. The model does not differentiate between products within the portfolio and assumes that each generates an average annual volume sales at an average annual price. This is a gross simplification which will be addressed in future versions of the model by differentiating between categories of products. Products are eventually discontinued, at a rate dependent on their rate of launch a product-lifetime ago.

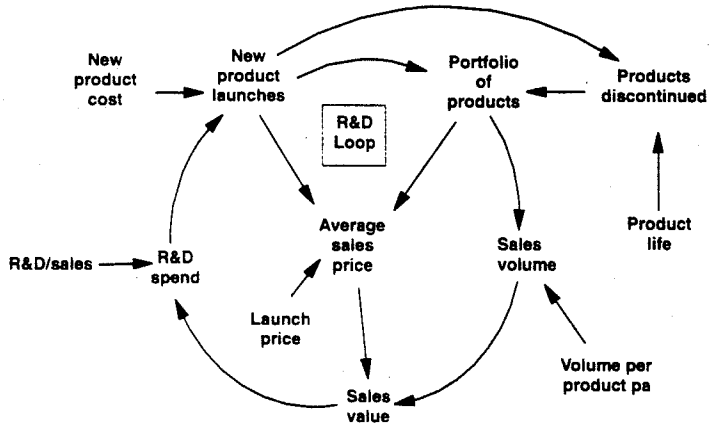

Figure 2. Core R&D Loop

The two volume-growth loops are shown in more detail in Figures 3 and 4. In both it is assumed that the industry has a long-term target growth rate which can be for either sales volume or sales value. Figure 3 shows schematically the first of these volume-growth loops and captures the effect of marketing effort (in excess of that needed to maintain product usage at current levels) on the average volume of drugs prescribed per patient per product per year.

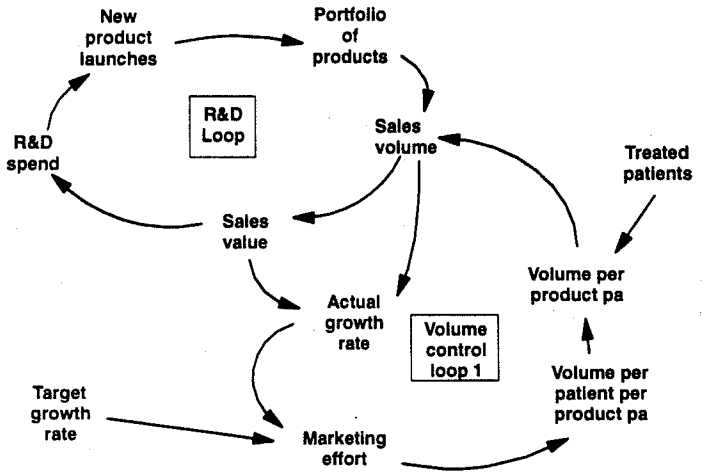

Figure 3. Volume growth loop 1

Marketing effort is measured in money terms and represents the costs of the sales force, promotional literature etc. which persuade the medical profession to prescribe the product to all relevant patients and for all relevant illnesses in a range of appropriate formulations. The concept of *treated* patients has been introduced to capture the fact that only a proportion of the population is ill at any particular time (measured in terms of disease prevalence) and, of this, a fraction only will have access to the full portfolio of drugs. The effects of aging can be captured by increasing disease prevalence.

The second volume-growth loop (Fig. 4) describes the other consequence of marketing expenditure - to bring about the expansion of product availability to geographic areas outside the country or region of original launch. Pharmaceutical companies rarely achieve simultaneous worldwide launches because of variations in registration timescales or, for many, because the upfront costs of concurrent global development and launch are prohibitive. It is usually a process of progressive territorial extension.

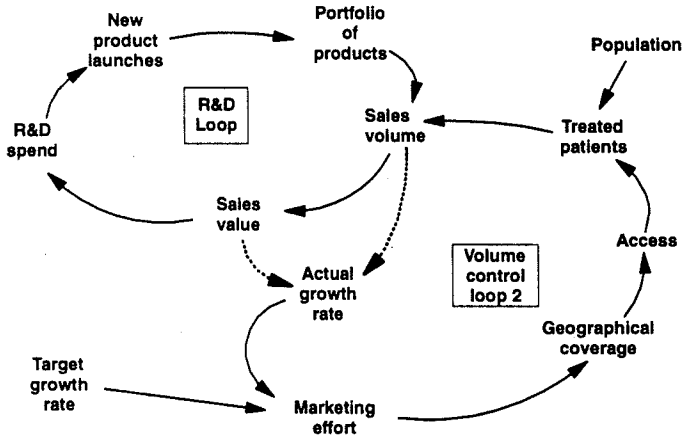

Figure 4. Volume growth loop 2

In the regional version of the model these loops are replicated (with different parameters) and contribute volume and value sales to the pharmaceutical companies which comprise the global industry. It is assumed that R&D is carried out on a global rather than regional or national basis.

2.3 Cost containment

As already described in the introduction, the organisations that fund pharmaceutical purchases have a wide range of methods available to them for controlling costs. These methods vary from country to country (particularly between the USA, where health-care is predominantly in the private sector, and many other countries, where public provision is the norm) but fall broadly into two categories. The first includes a variety of methods for controlling prices. The second comprises a portfolio of policies for limiting volume. In the multi-regional model these control loops are replicated.

Conflict arises because of the disparity between the annual drugs bill with which the funders are confronted and the budget that has been made available to them either by governments or, in the USA, by the aggregate health-insurance premiums from individuals and companies. The absolute size of these budgets has been determined by history but their growth rates are now increasingly seen in the context of overall economic growth. It is therefore a reasonable assumption that future health-care budgets will be tied closely to the rate of growth of GDP. It is also assumed, at least in the first instance, that the share of health-care expenditure accounted for by pharmaceuticals will be held broadly constant.

In both control loops (Figs. 5 and 6) the divergence between budget and actual expenditure is translated, with some allowance for the inevitable delays, into a desired expenditure. In the first loop this is then converted into a price adjustment mechanism which operates on the *average* sales price in the region. Some additional control is also imposed on launch prices.

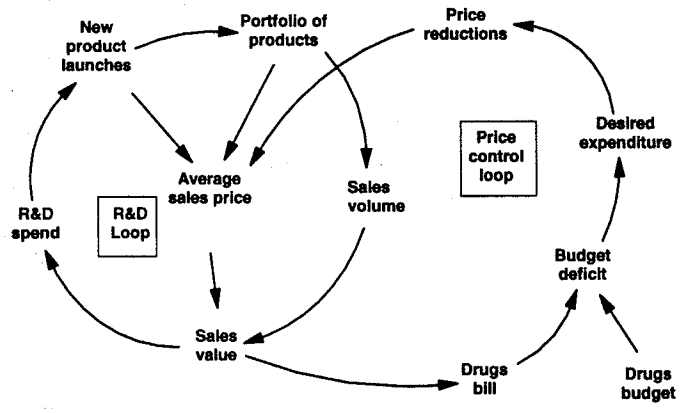

Figure 5. Price control loop

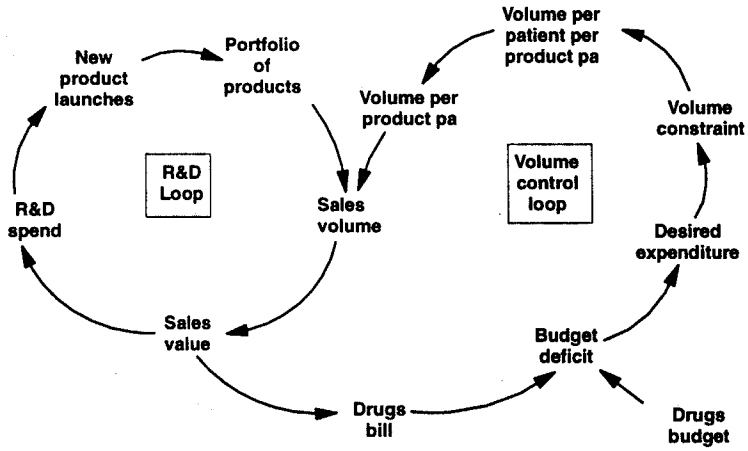

Figure 6. Volume control loop

In the second loop the desired expenditure is converted into an adjustment mechanism which controls the volume per patient per product pa (the average annual dose) on the assumption that the funders cannot directly influence the number of patients or the number of products in the portfolio. A more sophisticated representation might allow for a reduction in the number of products for which reimbursement is available, or for an increase in the percentage of the cost borne by the patients themselves and the deterrent effect this would have on people seeking treatment in the first place (ie indirectly limiting access).

It is not only funders who will experience financial pain. If the controls prove to be severe, either by intention or by accident, the pharmaceutical industry will also experience pain in the form of reduced profits. The model represents these effects, calculating profit by deducting R&D and marketing expenditures, production costs as well as un-attributable administrative costs from sales income. What happens to profits will influence the industry's policies on R&D expenditure, marketing and pricing (particularly of new products). In particular a cap is imposed on the proportion of sales revenue that can be spent on marketing and there is assumed to be a lower acceptable limit for profit margin at which R&D expenditure would be reduced to protect the remaining profits.

3 Validation of core assumptions

A set of core assumptions was chosen to ensure that the model broadly reproduced historical data over the decade 1980-1990. Data is sparse at this level of aggregation but estimates exist for total real sales value, ratio of R&D expenditure to sales, the number of products on the market, the number of new chemical entities and major line-extensions launched annually, the average R&D cost to the industry (and its rate of growth) for every successful product launched and the time taken to develop a product (**4**). From these it has been possible to estimate the values of variables for which past data are not directly available and to establish an internally consistent set of quantitative assumptions which constitute the historic foundation for the model.

A regional version of the model divides the global market into two - that in the Industrial World (N America, W Europe, Japan and Australasia) and that in the Developing World (the rest).

Since the models are seen as prototypes to explore a principle rather than to produce a quantitatively accurate representation of reality, only broad agreement has been sought between the model results and the historical data.

The key historical features replicated by the models are:

(a) A global market value of about $180bn in 1990, which has doubled in real terms (ie after allowing for general inflation) since 1980.
(b) An increase in real prices of 3-4% pa during the 1980s.
(c) An average of 11% of sales spent on R&D.
(d) A mean R&D expenditure per successful product that has increased over the preceding decade at a rate of about 7% pa in real terms to $360m in 1990.
(e) The launch of approximately 55 new products (new chemical entities plus major line extensions) in 1990, a figure that has declined from 60-65 in 1980.
(f) A total portfolio of some 3600 products currently prescribed worldwide, each with average sales of $50m pa.
(g) Sales in the developing world rising from 12% of the world market in 1980 to 16% in 1990.
(h) Average sales prices in the developing world substantially lower (20% assumed) than those in the industrial world due, at least in part to lower usage of more expensive drugs.

4 Scenarios

A wide range of scenarios describing likely states of the future market are being explored with the models. These range from a naive extrapolation of current trends to scenarios in which complex interactions are taking place between the industry and the controlling authorities.

4.1 Unrestrained growth

The "base case" against which all other scenarios can be compared assumes that the future will be similar to the 1980s, ie that the balance of cost constraint and growth aspirations will remain unchanged and that the growth of the pharmaceutical market will continue unabated. Sales in the industrial world continue to grow in real terms by about 6.5% pa and in the developing world by 9-10% pa. By 2020, the developing world's share of the market will be 30-35% and the global market will be growing at about 7.5% pa. This is likely to be 3-4% pa faster than world GDP, resulting in a pharmaceutical expenditure, in 2020, of some 2.5-3.5% of GDP compared to 1% today. The number of new products launched annually continues to decrease. If the pharmaceutical industry holds its unit costs of manufacture constant in real terms (as is assumed rather simplistically), by 2020 it would be earning a substantially higher profit margin than it does now. If other health-care costs also continue to rise at similar rates, this leads to a picture of the future with which many funding authorities would be uncomfortable.

The model can be used to examine the sources of the global sales growth. Of the 6.8% pa growth in the mid-1990s, 1.3% pa comes from population growth, 1.5% pa from increasing access (mainly in the developing world), 1.0% pa from increasing volume per capita (of which 0.4% pa is attributable to the growth of the number of available products) and 3.0% pa comes from increasing real prices.

4.2 Imposition of Price Control

Two versions of this scenario have been created: the first, in which the pharmaceutical industry accepts the control without response; the second, in which the industry responds in an attempt to maintain its profitability by raising marketing expenditure and hence promoting the growth of sales volume.

In both scenarios, sales growth is constrained such that each region's expenditure on drugs is pulled into line with a drugs budget that is, in its turn, constrained to grow at no more than a hypothetical GDP growth rate - assumed to be 3% pa in the industrial world (where controls are phased in over a 2-3 year period from 1990) and 7% pa in the developing world (control phased in from 2000). This is achieved by progressively reducing the average sales price of existing drugs. A limited degree of control is also imposed on the price-premium at which new products are launched.

In the non-reactive scenario the industry accepts the consequences of price control. After an initial transient during which the authorities impose price reductions to bring the situation under control and then relax in the light of the consequences, price rises in the industrial world are pulled back by about 2.5 percentage points, slightly more than offsetting the rate of price rise that has been imposed by the industry historically. The only price rise being permitted is that caused by adding new higher (but nevertheless constrained) price products to the portfolio. Similar consequences follow in the developing world, but the amount of long-term price control needed turns out to be smaller (about -1% pa) because the industrial world's restraint on launch prices carries over into the developing world and curbs sales growth there anyway. The effects on global average prices are shown, in comparison with the unrestrained scenario in Figure 7.

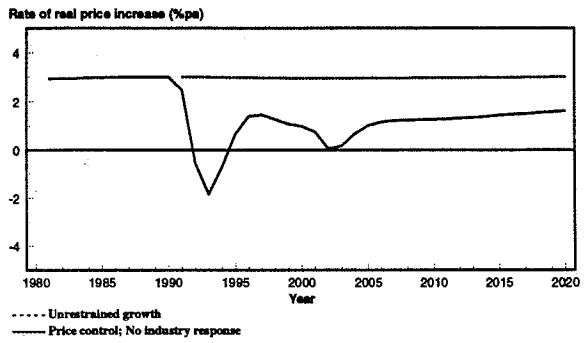

Figure 7. Effect of price control; no industry counter response

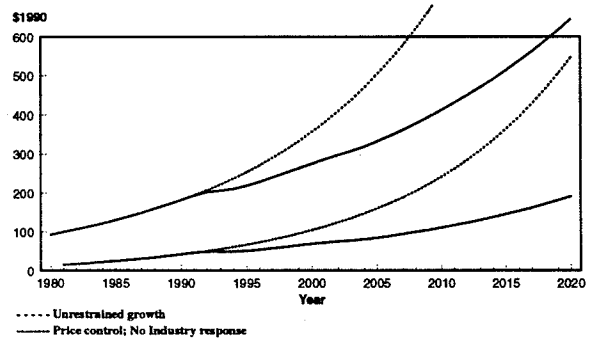

Figure 8. Industry sales and profit under price control; no industry counter response

The consequences for the industry's finances are serious but not devastating (Fig. 8). Sales and profits continue to grow, but at sharply reduced rates (for profits from 10% pa to a long-term 6% pa). The first round of price restraint, in the industrial world, stabilises profit margins at a (hypothetical) plateau at about 23-24% in the 1990s. The second round of price control holds the profit margin at 25% for about 5 years before it rises slowly towards 30% by 2020. The R&D:sales ratio is maintained throughout and the R&D spend therefore suffers the same diminution of growth as do sales. As a result the number of new products launched each year falls from 55 in 1990 to under 20 in 2020.

In the second scenario, the industry responds in both regions by raising the proportion of sales income devoted to marketing and drives up its volume growth (Fig. 9). The response in thedeveloping world has been chosen to be particularly marked (marketing:sales is capped at 25% in the Developed World and to 30% in the Developing). The authorities respond in turn by tightening the screw further on prices. They still achieve their budgetary targets but, this time by reducing long-term real price growth to -3% pa (Fig. 10). The effect on the industry's finances is much more dramatic. Profit margins collapse and profits fall in real terms from 2002 (Fig. 11). Although the industry is not forced into loss, from about 2010, because the profit margin has dropped below the critical level of 10%, the industry can no longer fund its R&D efforts; it reduces its R&D:sales ratio and the rate of launch of new products drops rapidly towards zero (Fig. 12).

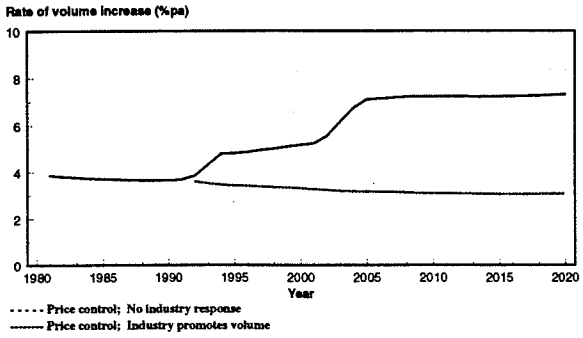

Figure 9. Effect of price control; Industry promotes volume growth

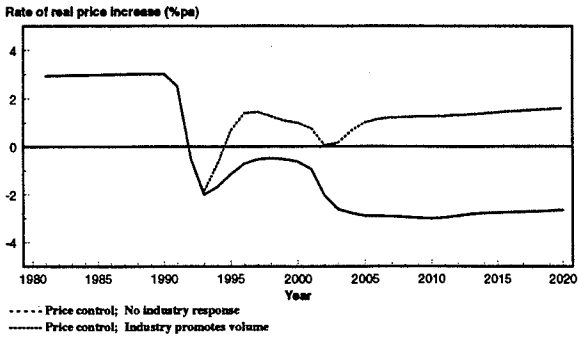

Figure 10. Effect of price control; Industry promotes volume growth

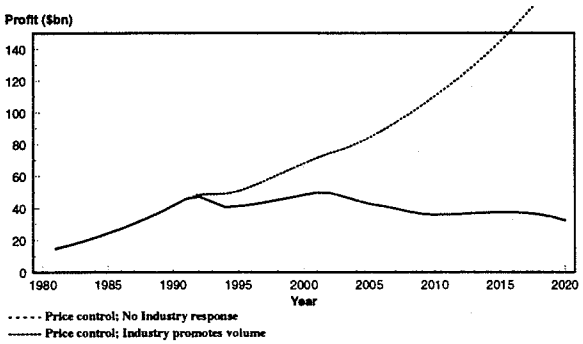

Figure 11. Industry profit under price control; industry promotes volume growth

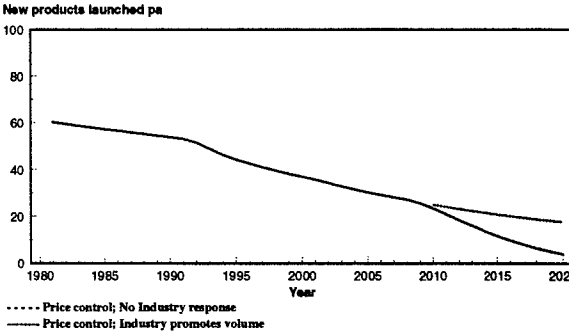

Figure 12. New product launches under price control; industr promotes volume growth

In both these scenarios the industry undergoes a dramatic long-term transformation. It becomes a bulk supplier of drugs which undertakes little R&D and creates few new products, ie a high-volume/low-price non-innovative industry. What is more, in the second, the industry compounds its problems through the adoption of inappropriate policies. It seems that no response would have been better than the wrong response.

4.3 Imposition of volume control

The same objectives of budget growth are ascribed to the controlling authorities as in the previous scenarios. However, in this case, the budgetary pressure is relieved by holding down volume growth and leaving prices to follow an unconstrained path. In the first scenario, the industry passively accepts the restraint. The impact on profit is similar to, but less severe than that under price control; (real) profit growth is suppressed during the five years or so that it takes for the new policies to get a grip, but then recovers, albeit to a rate some 3% pa below the pre-intervention level. The industry's ability to finance its R&D is not impaired but the lower sales growth causes a similar decline in the rate of launch of new products. The main effect of the controls is to reduce significantly the volume of drugs prescribed per person (by as much as 30% in the industrial world). This scenario, therefore, has social

implications in that in the long term the market is one characterised by high-price and low-volume and is therefore one in which access to treatment is being rationed and the cost of treatment is high for those who receive it.

The industry has two ways in which it can try to offset the effects of the controls. It can increase marketing expenditure in order to promote volume usage of its products - thereby entering into direct conflict with the authorities - or it can seek to raise prices. In the first case the (model) authorities just bear down harder still on volume, the industry achieves no higher growth but incurs the additional costs of the marketing effort, thereby further depressing its profitability. If it puts up prices, the industry does protect its profit margin to some degree. However, the authorities squeeze volume to such an extent that there must inevitably be a deterioration in the quality of health-care being provided - with the associated political risk to both the authorities and the industry.

4.4 Imposition of both price and volume control

Scenarios have been generated in which the authorities impose a mix of price and volume controls and the industry responds with its own mix of pricing and marketing policies. Different outcomes for the industry and for society at large can be obtained depending on the balance between the authorities' price and volume policies and the industry's choice of response. In broad terms, these compromise futures lie between the extremes described in sections 4.2 and 4.3 above. It is in the creation of such scenarios that real learning starts to take place and the model begins to reward the investment that has been made.

5 Conclusions

The prototype models have demonstrated that it *is* possible to create a simplified but realistic representation of the global pharmaceutical industry and the interactions between its growth aspirations and the authorities' cost containment objectives. Although not too much weight should be put on the preliminary quantitative results described above, the exercise has shown the likely consequences of the world continuing along the path taken in the past decade and can explore the effects on the industry of different types and strengths of cost containment policies. It can demonstrate not only the direct consequences of these constraints on the industry, but also the industry's possible reactions and thus the ultimate effectiveness of both the authorities' and the industry's strategies. The complex feedback nature of the total system will often lead to unexpected or counter-intuitive outcomes.

The models have not yet been exhaustively explored. They contain control and response mechanisms yet to be examined, eg increased patient co-payments and their consequences in terms of restricting access. There are many more that could and should be incorporated eg the pressure to use generic rather than branded drugs.

An important area lying outside the current boundaries of the model is the interaction between drug consumption and other forms of medical intervention. The industry has always argued that ill-considered reductions in expenditure on drugs will frequently result in increased expenditure elsewhere in the health-care system.

Currently the models are non-specific in terms of the issues that they address. In this form, they could be developed as a "microworld" - perhaps as a learning tool for senior management or as an aid to constructive dialogue between the industry and the authorities. They could also be developed to provide the conceptual and quantitative underpinning of a scenario planning process. More productively, the approach could be developed in a more focused, perhaps country-specific way, in the context of a clearly defined issue or problem.

6 References

1. Dower, M. 1993. A year to retrench, restructure and rethink. *Scrip Magazine* January 1994: 2-5.
2. de Geus, A.P. 1992. Modelling to predict or learn? *European Journal of Operational Research* 59: 1-5.
3. Morecroft, J.D.W. 1992. Executive knowledge, models and learning. *European Journal of Operational Research* 59: 9-27.
4. Derived from a variety of published and unpublished pharmaceutical industry sources.